# Can co-speech gestures alone carry the mental time line?

Winter, Bodo; Duffy, Sarah

*DOI:*
10.1037/xlm0000836



*Document Version*
Peer reviewed version

*Citation for published version (Harvard):*
Winter, B & Duffy, S 2020, 'Can co-speech gestures alone carry the mental time line?', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 46, no. 9, pp. 1768-1781.
https://doi.org/10.1037/xlm0000836

# Can co-speech gestures alone carry the mental time line?

Bodo Winter[1] & Sarah E. Duffy[2]

[1] University of Birmingham

Dept. of English Language and Applied Linguistics

[2] Northumbria University

Dept. of Humanities

**Author note:** The full data and code are accessible online: https://osf.io/vc8bw/

(OSF Repository DOI: 10.17605/OSF.IO/VC8BW )

**Abstract**

Time and space have been shown to be interlinked in people's minds. To what extent can co-speech gestures influence thinking about time, over and above spoken language? In this study, we use the ambiguous question "Next Wednesday's meeting has been moved forward two days, what day is it on now?" to show that people either respond "Monday" or "Friday," depending on gesture. We manipulated both language (using either the adverb "forward", or the adverb "backward") and gesture (forward and backward movement), thus creating matches and mismatches between speech and gesture. Results show that the speech manipulation exerts a stronger influence on people's temporal perspectives than

gesture. Moreover, the effect of gesture disappears completely for certain hand shapes and if non-movement language is used ("changed by two days" as opposed to "moved by two days"). We additionally find that the strength of the gesture effect is moderated by likeability: when people like the gesturer, they are more prone to assuming their perspective, which completely changes the meaning of forward and backward gestural movements. Altogether, our results suggest that gesture does play a role in thinking about time, but this role is auxiliary when compared to speech, and the degree to which gesture matters depends on one's social relation to the gesturer.

**Key words:** perspective; metaphor; ambiguity; moving time; moving ego; temporal deixis

## 1. Introduction

When people talk about time, they often gesture (Casasanto & Jasmin, 2012; Núñez, Cooperrider, Doan, & Wassmann, 2012; Núñez & Sweetser, 2006). For example, an American English speaker may move their hand forward and to the right when using the expression "looking to the future" (Walker & Cooperrider, 2016). These observations from spontaneous co-speech gestures correspond to experimental research which demonstrates that people think about time both in terms of horizontal (transversal) and sagittal (front-back) space (Hartmann & Mast, 2012; Miles et al., 2010). For example, participants from cultures in which the direction of orthography is left-to-right respond more quickly to past events with their left hand and to future events with their right hand (Ouellet et al., 2010; Weger & Pratt, 2008). In addition, decades' worth of linguistic research has shown that speakers of English and other languages use spatial terms to talk about time (Clark, 1973; Haspelmath, 1997; Lakoff & Johnson, 1980; Moore, 2014), such as when saying that "Halloween is around the corner", "your past is behind you", or that "spring has passed" (examples from the GloWBE corpus, Davies, 2009)." Together, these different strands of research suggest a deep cognitive connection between space and time (Bender & Beller, 2014; Bonato et al., 2012; Casasanto & Boroditsky, 2008; Winter et al., 2015).

Linguistic research distinguishes between different types of time metaphors. "Deictic" metaphors involve the positioning of events relative to the ego (Clark, 1973; Lakoff & Johnson, 1999), "sequential" metaphors situate events in relation to

one another, as part of a sequence (Moore, 2006; Núñez, Motz, & Teuscher, 2006), and "absolute" metaphors fix events in relation to the forward-flowing trajectory of time (Evans, 2013; Kranjec, 2006). Particular attention has been directed toward the distinction between "Moving Ego" and "Moving Time" perspectives in the domain of deictic time metaphors. Whereas Moving Ego metaphors depict time as a static landscape, across which the active ego moves ("We are approaching the deadline"), Moving Time metaphors construe time as a conveyor belt on which events move, relative to a stationary ego ("The deadline is approaching") (Clark, 1973; Lakoff & Johnson, 1999; McGlone & Harding, 1998).

In this paper, we use the contrast between Moving Time and Moving Ego perspectives to study the interconnection between language and co-speech gesture. Specifically, we asked participants the question "Next Wednesday's meeting has been moved forward two days, what day is it on now?," a question that is known to be ambiguous (McGlone & Harding, 1998). A recent meta-analysis has shown that without any further context, about 54% of all English speakers respond "Friday," and 46% respond "Monday" (Stickles & Lewis, 2018). This ambiguity has spurred a number of studies to explore possible factors contributing to the interpretation of the question (Boroditsky & Ramscar, 2002; Duffy et al., 2014; Duffy & Evans, 2017; Duffy & Feist, 2014; Matlock et al., 2011; Ramscar et al., 2010). In general, factors that involve self-motion (in line with the Moving Time construal) lead to more "Friday" Responses, as if the meeting had been moved from Wednesday to further away in time. In contrast, factors that involve motion toward oneself (in line with the Moving

Time construal), lead to more "Monday" responses, as if the meeting had been moved from Wednesday closer in time (Boroditsky & Ramscar, 2002).

One of the factors that affects people's propensity to adopt a Moving Time or a Moving Ego perspective is gesture. Specifically, it has been found that if the speaker is standing next to the addressee so that both the speaker and the addressee share the same perspective, gesturing *forward* away from the torso leads to more "Friday" responses, and gesturing from an extended position *backward* toward the torso leads to more "Monday" responses (Jamalian & Tversky, 2012; Lewis & Stickles, 2017). In addition, Lewis and Stickles (2017) (Experiment 1) found no statistically reliable interaction between the factor perspective (standing next to the participant versus standing opposite of them) and the directionality of the hand movements. Specifically, when a speaker is standing opposite of the addressee (face-to-face), a movement away from the speaker is a movement toward the addressee. Thus, the fact that Lewis and Stickles (2017) find more "Friday" with movements away from the speaker (toward the addressee), suggests that addressees assume the speaker's perspective, an other-centric rather than ego-centric perspective.

Other researchers have looked at linguistic factors influencing people's responses to the ambiguous question (Elvevåg et al., 2011; Loermans et al., 2019; Stocker & Hartmann, 2019). Of particular interest to this study is the fact that Feist and Duffy (2015) found that both the spatial adverb ("backward" versus "forward") and the verb (e.g., "rushed" versus "moved") change response behavior. Certain verbs are strongly biased toward a Moving Time perspective, such as the verb

"rush," which consistently lead to more "Monday" responses; other verbs are strongly biased toward a Moving Ego perspective, such as the verb "carry," which consistently lead to more "Friday" responses. In addition to this, Feist and Duffy (2015) found that the adverb "forward" led to more Monday responses than its directional opposite, "backward."

Here, we use the fact that certain linguistic expressions are biased toward "Monday" or "Friday" to look at the interplay of language and gesture. Specifically, we ask the question: What happens if there are mismatches between the directionality implied by language ("forward/backward") and the directionality implied by the gesture? If, for example, the speaker uses the adverb "backward" but moves the hand forward, do people respond "Friday," thus following the gesture, or do they respond "Monday," thus following the spoken language? In addition, we sought to investigate whether gesture can carry the time line alone, that is, what happens when forward and backward movements are paired with linguistic expressions that are non-directional ("moved by two days")? And what happens when gestural movement is combined with language that makes no explicit reference to movement whatsoever ("changed by two days")?

To the extent that we are interested in showing that gesture changes the interpretation of the Next Wednesday question, we aim to replicate Jamalian and Tversky (2012) and Lewis and Stickles (2017). However, moving beyond these established findings, we use the question to look at the interplay of language and speech, and how both together can change people's thinking about abstract

concepts, such as time. Our experiments are relevant above and beyond our understanding of temporal language, as they assess the role of gesture in the interpretation of spoken messages more generally. It is clear that in some situations, the interpretation of a spoken utterance *completely* depends on gesture, such as when a speaker says "it was this long" while moving the hands apart. Without the gesture demarcating a specific length, it is not clear what "this long" means. Similarly, expressions such as "look over there" make no sense without concomitant pointing gestures. More generally, gesture researchers are keen to emphasize that linguistic communication is intrinsically multimodal (Goldin-Meadow, 2005; Kendon, 2004, 2014; McNeill, 1992; Müller, 2009; Streeck, 2009), involving *both* speech and gesture. However, to what extent does communication depend on speech, and to what extent does communication depend on gesture alone?

### 1.1. Overview of experiments

In Experiment 1, we attempted to replicate the finding, reported by Lewis and Stickles (2017), that in the opposing view (the addressee looking straight at the face of the speaker), gestures away from the speaker lead to more "Friday" responses, compared to gestures toward the speaker, which lead to more "Monday" responses. In addition, following Feist and Duffy (2015), we investigated what happens when contrasting "forward" with "backward" adverbial expressions. While we successfully found an effect of spatial adverb ("backward" expressions lead to more

"Monday" responses than "Friday" responses), thus replicating Feist and Duffy (2015), we failed to find any effect of gestural direction.

The absence of a gestural direction effect was unexpected. We thus explored two potential changes to the experiment: First, Experiment 1 used a different hand shape from previous studies, so we explored two alternative hand shapes in Experiment 2, including a hand shape that was previously used in Jamalian and Tversky (2012) and Lewis and Stickles (2017). In addition, post-experiment debriefing of Experiment 1 suggested that people had strong positive or negative opinions about the depicted speaker seen in the video (the first author of the study). We reasoned that if participants did not like the speaker, they would be less likely to assume his perspective, which would completely alter the response pattern. In particular, if the hands are moving away from the speaker and toward the addressee, this would lead to a Moving Time conceptualization ("Monday") if participants did not adopt the perspective of the speaker. An effect of gestural direction could be masked in the average if some people adopt the speaker's perspective, and others do not. To explore this, Experiment 2 asked participants to what extent they liked the speaker.

Using the altered hand shapes, Experiment 2 found a gestural direction effect. In addition, we found a weak effect of speaker likeability which showed that, indeed, participants were more likely to respond "Friday" for gestures moving away from the speaker when they liked the speaker more strongly. All further experiments thus included the likeability scale.

Equipped with a gesture that we knew could produce a directional effect, Experiment 3 revisited the original goal of creating matches and mismatches between speech and gesture. We replicated the directional effect observed in Experiment 2, as well as the effect of adverb observed in Experiment 1 ("forward" versus "backward"). However, the verbal effect was much stronger than the directional effect, even though we used very pronounced two-handed gestures. In addition, Experiment 3 replicated the effect of speaker likeability, thus providing further evidence for the idea that social factors moderate perspective taking in this task.

Finally, Experiment 4 changed the words "moved forwards/backwards by two days" to simply "moved by two days" (without spatial adverb) or "changed by two days" (without movement verb). There was a weak directional effect for the "move" verb, but no directional effect whatsoever for the "change" verb, which suggests that gesture only influences temporal perspectives when the concomitant speech implies movement. In addition, Experiment 4 provided additional evidence for a weak effect of likeability.

In a final analysis, we combined the data from all four experiments to try to resolve a puzzle that emerged in this series of experiments, namely, the overall proportion of "Monday" and "Friday" responses shifted quite considerably as a function of what day of the week the participants were tested on. These additional analyses point to important methodological issues that need to be controlled for in future studies using the "Next Wednesday's meeting" question.

## 2. Experiment 1

### 2.1. Methods

#### 2.1.1. Procedure and stimuli

We used Qualtrics to collect responses online via Amazon Mechanical Turk. After consenting, participants read the following instructions:

> "On the next screen, you will see a video of a person.
>
> Focus on the screen. Please play the video only once.
>
> Listen closely to what the person says!
>
> Do not spend too much time about answering the question, try to respond as fast as possible. Do not change your answer—we are interested in your initial reaction and a quick, intuitive judgment."

The participant then saw a video where the speaker (the first author) said "Next Wednesday's meeting has been moved ***forward*** / ***backward*** two days. What day has the meeting been rescheduled to?" Sentence stress was put on the directional preposition, "forward" in one condition and "backward" in the other. Orthogonal to this linguistic manipulation the speaker either moved the palm (facing upward) from about chest height positioned just above the stomach to a position further away from the body (forward gesture condition), or the reverse

(backward gesture condition), as depicted in Figure 1a. We thus adopted a 2 X 2 design, with the factors "Gesture Direction" and "Language Direction."

The palm-up-open-hand (PUOH) gesture was chosen because this gesture is commonly associated with the presentation of ideas (Cooperrider, Abner, & Goldin-Meadow, 2018; Kendon, 2004, Ch. 13; Müller, 2004), and it has been claimed to be very frequent in discourse (Mittelberg, 2017; Parrill, 2008). We reasoned that a PUOH gesture would be less marked and thus more ecologically valid than the gesture used by Jamalian and Tversky (2012) and Lewis and Stickles (2017), which was a flat hand shape with the palm facing inward (toward the speaker), as shown in Figure 2a (left). In addition, a PUOH gesture may further suggest a certain openness of the speaker and a willingness to convey information to the listener.

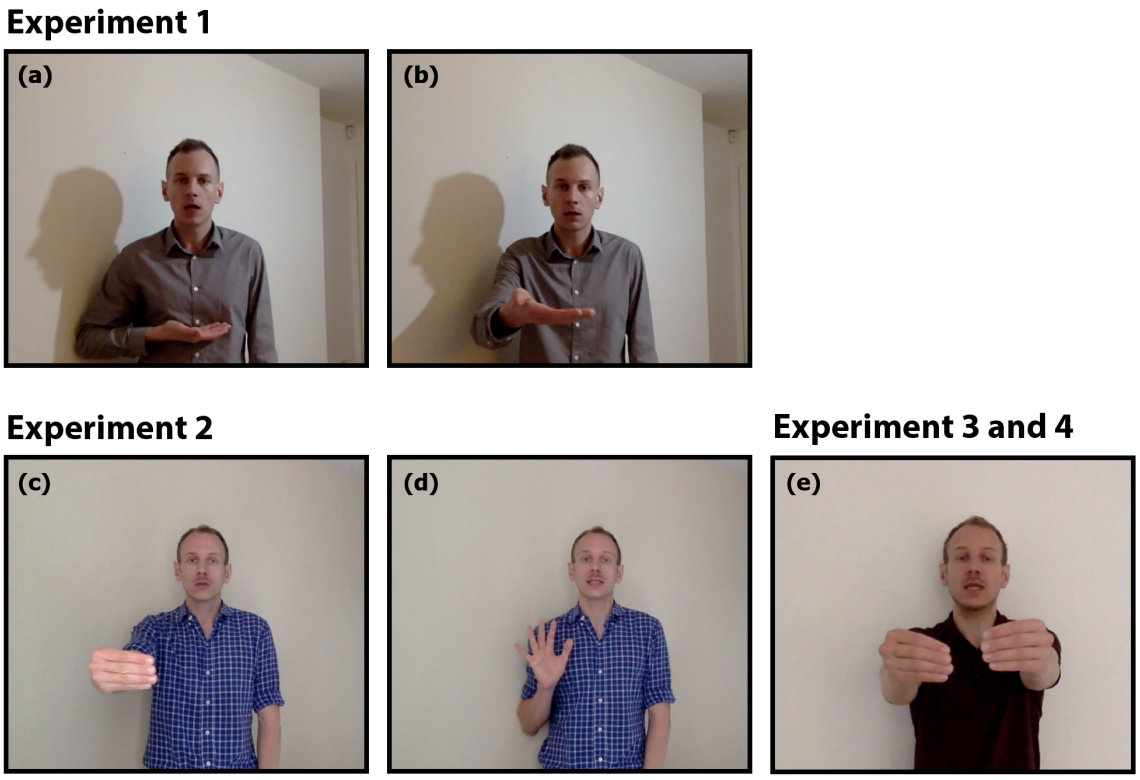

**Figure 1.** (a) Start position and (b) end position of the forward-moving gesture in Experiment 1, with in a palm-up configuration; single-handed versions of the (c) flat "wall" and (d) open hand "push" gestures used in Experiment 2; (e) the two-handed flat "wall" gesture used in Experiments 3 and 4

Participants then had to click onto the NEXT arrow to move to the next screen. Where we asked the question "what day of the week has the meeting been rescheduled to?", with an open-ended text entry box immediately below the question.

After participants provided their primary response, we asked several control questions, such as "Were you able to view the video completely?" (Yes — Maybe — No), and "If you were able to view the video, how many times did you view it?" (open ended text entry), "Did you notice the person's gestures?" (five-point scale from "Definitely Yes" to "Definitely not"), and finally, an open-ended question phrased as follows: "Any observations, thoughts, ideas you want to share with us about the person you just saw or the question you were being asked?"

We also asked a comprehension question ("What is 9 + 4?") to make sure that participants were paying attention. Finally, to complete the survey, participants had to answer several demographics questions (age, gender, handedness, and self-reported native languages).

This study was approved by the University of Birmingham Research Humanities and Social Sciences Ethical Review Committee (ERN_17–0040). Consent

was obtained electronically. The Qualtrics survey file (.qsf) and the video files (.mp4) are available under the following publicly accessible Open Science Framework repository (Winter & Duffy, 2020, OSF DOI: 10.17605/OSF.IO/VC8BW):

https://osf.io/vc8bw/

### 2.1.2. Participants

239 participants were recruited via Amazon Mechanical Turk for a 0.30 USD reimbursement. Although not without its own share of problems, Amazon Mechanical Turk is known to be a useful source for collecting behavioral data online (J. Bohannon, 2011; Paolacci et al., 2010; Paolacci & Chandler, 2014; Rouse, 2015), including linguistic data (Sprouse, 2011). Moreover, Amazon Mechanical Turk has been used to great effect in the context of the "Next Wednesday's meeting" question (Lewis & Stickles, 2017; Stickles & Lewis, 2018). Only Turkers who were located in the United States of America and had an overall approval rate of at least 92% were able to view our task and allowed to participate in it. We used TurkGate (Goldin & Darlow, 2013) to make sure that participants who participated in this experiment could not partake in any of the future experiments.

All participants responded to the control question (9 + 4 = ?) correctly. We excluded 12 participants who reported to be non-native speakers and/or who reported to experience viewing issues in our post-experiment questionnaire (e.g., buffering issues half-way through the video). We excluded an additional 36

participants who did not provide either “Monday” or “Friday” responses, including such responses as “Wednesday”, “Saturday”, “Monday or Tuesday,” “two days from the original date,” or “They didn’t say. Just that it would be moved forward 2 days.” Thus, a total of 48 participants (20%) were excluded from the analysis. The final sample consisted of 126 men and 65 women (average age 34, range 19 to 67).

### 2.1.3. Statistical analyses

All statistical analyses were conducted with R version 3.3.1 (R Core Team, 2016) and the packages “stringr” (Wickham, 2016a) and “tidyverse” (Wickham, 2016b) for data processing. All statistical analysis code and data is made available under the OSF repository associated with this publication: https://osf.io/vc8bw/

We used the package “brms” 2.9.0 (Bürkner, 2017) for analyze the binary response variable “Monday” versus “Friday” with a Bayesian logistic regression model. We included the predictors Gesture Direction (forward versus backward) and Language Direction (“forward” versus “backward”). In the analysis of this experiment and all following experiments, all categorical predictors were deviation-coded (-0.5 = backward language or backward gesture, +0.5 = forward language or forward gesture) to facilitate the interpretation of main effects in the presence of interactions. In all following models, the reference level of the dependent variable was set to “Monday,” thus, the coefficients below report the changes in the odds of observing a “Friday” response.

We used default priors for the intercept and the standard deviation as well as weakly informative priors with normal distributions centered at zero (*SD* = 1) as priors for the coefficients. Hamiltonian MCMC sampling was conducted with 8 chains and 8,000 iterations (2,000 of which were warm up), resulting in a total of 48,000 posterior samples used for inference. The same priors and MCMC specifications were used for all of the following experiments. There was no indication of convergence issues (all Rhat values = 1.00, no divergent transitions) in the model for this experiment or all following experiments.

**2.2. Results**

Figure 2 displays the proportion of "Monday" and "Friday" response as a function of Language Direction and Gesture Direction. Overall, 136 participants responded "Monday" (71 %), compared to 55 participants who responded "Friday" (29 %). When broken up by the Language Direction factor, 83 participants in the "backward" adverb condition responded "Monday" (86%), compared to only 14 who responded "Friday" (14%). In contrast, responses were more mixed in the "forward" adverb condition, with 54 participants responding "Monday" (56%) and 41 participants responding "Friday" (44%). The Bayesian logistic regression model indicated a positive coefficient for the Language Direction factor (log odd increase of "Friday" responses in the "forward" condition = 1.39), with a 95% credible interval far away from zero [0.74, 2.06]. In fact, not a single posterior sample for this

coefficient was below zero, indicating very strong evidence for an effect of the adverbial manipulation, Pr ($\hat{\beta} > 0 = 1.0$).

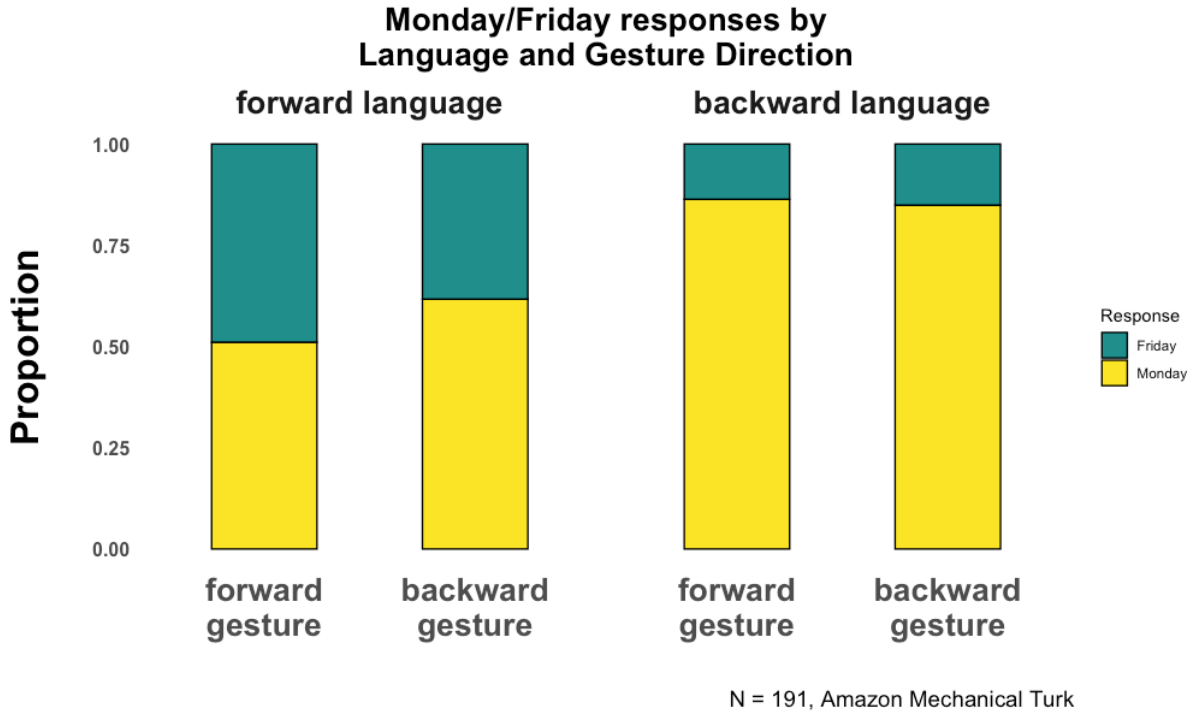


**Figure 2.** Proportions of "Friday" responses (green) and "Monday" responses (yellow) differ by adverb (more "Friday" responses for "forward") but not by the direction of the gesture

What about the effect of Gesture Direction? When the gesturing hand was moved forward (away from the speaker), 62 participants responded "Monday" (68%) and 29 participants responded "Friday" (29%). There was little difference to the condition where the gesturing hand moved backward (toward the speaker), where 74 participants responded "Monday" (74%), compared to 26 participants who responded "Friday" (26%). The Bayesian logistic regression model estimated the

coefficient of the Gesture Direction effect to be very small ($\hat{\beta} = 0.18$) with a wide 95% credible interval that included zero [-0.47, 0.82]. The estimated posterior probability of this effect being above zero, Pr ($\hat{\beta} > 0$), was 0.70. There similarly was no compelling evidence for an interaction between Language Direction and Gesture Direction, with the coefficient of 0.37 (more "Friday" responses in the "forward" language and the forward gesture condition) having a wide credible interval [-0.75,1.51]. The posterior probability of the effect being above zero was 0.74.

### 2.3. Discussion

Experiment 1 failed to find compelling evidence for an effect of Gesture Direction. There was, however, a very reliable effect of Language Direction. Somewhat puzzingly, this effect goes in the *opposite* direction of what is reported in Feist and Duffy (2015). Whereas they reported more "Monday" responses for the "forward" adverb, compared to "backward," the present study found that the "forward" adverb increased the proportion of "Friday" responses. Similar to Feist and Duffy (2015), we found an overall bias toward "Monday" responses. This deviates from most studies on the "Next Wednesday's meeting" question, which consistently find a small but reliable overall bias toward "Friday" responses (Stickles & Lewis, 2018). Our meta-analysis (see section 6) will shed light on why we observed overall more "Monday" responses compared to other studies, however, we currently have no compelling explanation as to why the "forward/backward" manipulation lead to the opposite effect of that reported in Feist and Duffy (2015). One possibility is that the

meanings of these spatial adverbs is subtly different in the UK population tested in Feist and Duffy (2015), compared to the present population, which was US-based. This is a likely explanation, especially since dialectal differences in the interpretation of the Next Wednesday question have already been noted for other languages (Stocker & Hartmann, 2019).

In contrast to Jamalian and Tversky (2012) and Lewis and Stickles (2017), we found no strong evidence for a reliable effect of Gesture Direction. This is unlikely due to a difference in statistical power between these experiments: Experiment 1 included more participants than Experiment 1 in Lewis and Stickles (2017) ($N$ = 168) and many more than the experiment conducted by Jamalian and Tversky (2012) ($N$ = 40).

Could it be that the lack of a Gesture Direction effect has to do with the fact that we used Amazon Mechanical Turk, or that people simply did not pay attention to the screen, thus not noticing the gestures? This seems unlikely. First, Lewis and Stickles (2017) found reliable gesture effects using Amazon Mechanical Turk. Second, in our post-experiment debriefing question, we asked participants "Did you notice the person's gestures?", 176 participants (92%) responded "Definitely yes," and furthermore 11 participants (6%) responded "Probably yes." In contrast, only 4 participants in total (<3%) responded "Might or might not," "Probably not," or "Definitely not" to this question. Thus, lack of attention to gesture is unlikely the cause of our failure to obtain a gesture effect.

The most likely explanation behind the null result for Gesture Direction is the choice of hand configuration. Lewis and Stickles (2017) also used an open handed gesture, but with the palm vertically aligned, facing toward the torso of the speaker, similar to what is described in Jamalian and Tversky (2012) (see Figure 2c). It is quite likely that the palm-up open hand gesture (PUOH) we used was interpreted in a different, non-temporal manner. In particular, consistent with the meaning of this hand configuration reported in the literature (Cooperrider et al., 2018; Kendon, 2004; McNeill, 1992; Mittelberg, 2017; Müller, 2004; Parrill, 2008), this gesture may have been interpreted as merely "presenting information" rather than indicating a movement along the time line. In Experiment 2, we manipulated the form of the gesture to see whether this would yield any directional effects.

An additional explanation is suggested by the participants' responses to the debriefing question. Some had very positive impressions of the speaker saying such things as "He was very articulate" or "He seemed like a nice guy." Others noted that the speaker seemed strange or foreign, such as "He isn't native to the US, and uses odd hand gestures," "Is he from a different planet?," and "He seems like he wants to be a mentalist." We thought that perceived likeability of the speaker may be a moderating factor in this experiment since the directional effect of the gesture depends on whether one is willing to assume the speaker's perspective or not. Thus, Experiment 2 included two socially relevant scales to investigate this phenomenon.

## 3. Experiment 2

### 3.1. Methods

The main purpose of Experiment 2 was to assess whether the Gesture Direction effect can be obtained at all if the execution of the gesture is changed. We sought to replicate the Gesture Direction effect reported in Lewis and Stickles (2017), which used a "flat" palm gesture, as shown in Figure 2c. In addition, we explored one additional hand configuration, the "pushing" gesture shown in Figure 2d. We included both single-handed and two-handed versions of both of these gestures. Whereas Lewis and Stickles (2017) used single-handed gestures, Jamalian and Tversky (2012) used two-handed gestures, which may be more salient to the addressee. As Experiment 2 was focused on exploring the role of gestural execution, we did not manipulate the language as in Experiment 1, instead keeping with the more canonical "forward" form of the question. Thus, the design of Experiment was 2 X 2 X 2 (Gesture Direction * Hand Shape * Hand Number).

A few subtle cosmetic changes were undertaken to improve Experiment 1 further. First, rather than asking "What day has the meeting been rescheduled to?", we asked "what day is it on now?", which more closely follows Lewis and Stickles (2017). Second, in Experiment 1, the response box appeared on a separate screen (after participants clicked ">>>"), which perhaps made responses less immediate, and it furthermore introduced an additional (rightward-pointing) spatial element. This time, we presented the text entry box on the same screen immediately beneath the video. Third, we changed the video so that it included improved lighting conditions and a clearer background. The new video also included a 2 second intro

for which the words "Watch this video" appeared on a black screen. This was included to ensure that participants paid enough attention to the video.

Following on from Experiment 1, where we had to exclude many participants due to not responding either "Monday" or "Friday," and after the null result for Gesture Direction in Experiment 1, we decided to run the new experiment with a slightly increased sample size ($N$ = 294). Unfortunately, we ended up with a similar sample size to Experiment 1 ($N$ = 199) due to an even larger number of exclusions (32% exclusions overall due video issues, responses that were neither "Monday" nor "Friday," etc.). The final sample included 125 men and 74 women (average age = 36, range 18 to 72).

In correspondence with the 2 X 2 X 2 design, the Bayesian logistic regression model included the factors Gesture Direction (backward versus forward), Hand Shape (flat versus push), and Hand Number (single-handed versus two-handed), all deviation coded (-0.5, +0.5).

To explore social factors in influencing perspective taking, we added four items from Reysen's likeability scale (Reysen, 2005), asking whether participants thought that the "person in the video" is "warm," "approachable," "friendly," and "likeable." In addition, we added the two best-performing social questions from an abbreviated Autism Quotient (AQ) measure (Allison et al., 2012): "I find it easy to work out what someone is thinking or feeling just by looking at their face" and "I find it difficult to work out people's intentions." For both the likeability scale and

AQ measure, we used a five point scale ranging from 1 ("strongly agree") to 5 ("strongly disagree").

As the inclusion of likeability and AQ question increased the length of the experiment, we raised the fee paid to Amazon Mechanical Turk participants from 0.30 to 0.40 USD.

### 3.2. Results

Figure 3 shows the results of Experiment 2. Overall, there were slightly more participants responding "Friday" (*N* = 112, 56%) than "Monday" (*N* = 87, 44%). This time, there was a reliable effect of Gesture Direction. Averaging across all forms of gestural execution, forward movements led to more "Friday" responses (*N* = 78, 76%) than "Monday" responses (*N* = 24, 24%). In contrast, backward movements led to relatively more "Monday" responses (*N* = 63, 65%) than "Friday" responses (*N* = 34, 35%). The Bayesian logistic regression indicated a strong effect of Gesture Direction ($\hat{\beta} = 1.76, 95\%\ CI[1.16, 2.38]$), with the posterior probability of this main effect being larger than zero $\Pr(\hat{\beta} > 0) = 1.0$ (no single posterior sample below zero).

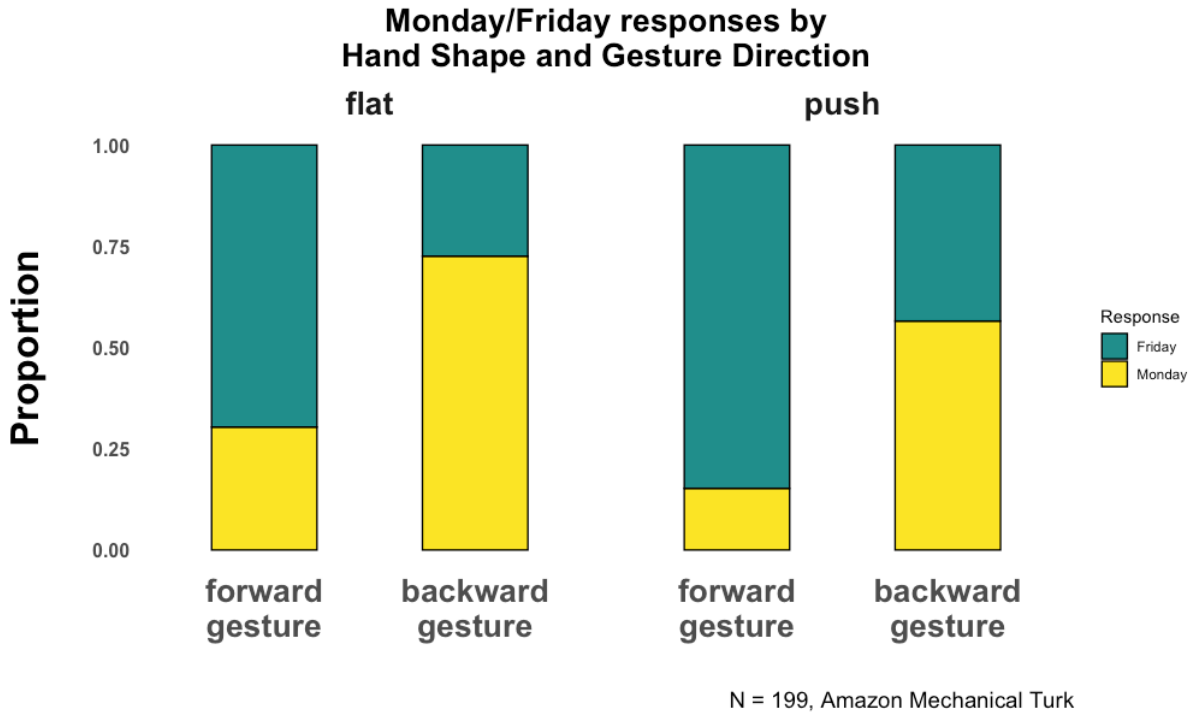


**Figure 3.** Gesture direction (forward/backward) reliably shifted the proportion of "Monday" and "Friday" responses regardless of whether a "flat" or "push" gesture was used; results for Hand Number not shown as they similarly did not produce any stark differences in the degree to which Gesture Direction influenced responses

There also was evidence for a main effect of Hand Shape ($\hat{\beta} = 0.71, 95\%\ CI[1.16,2.38], \Pr(\hat{\beta} > 0) = 0.99$), with the "push" gesture leading to more "Friday" responses ($N$ = 59, 64%) than "Monday" responses ($N$ = 33, 36%), compared to the "flat" gesture, which evoked about the same number of "Friday" responses ($N$ = 53, 50%) and "Monday" responses ($N$ = 54, 50%). There was no evidence for a reliable main effect of Hand Number, nor was there any indication for strong two-way or three-way interaction effects (all 95% credible intervals included zero). There was some mild evidence for a two-way interaction between Hand Number and

Gesture Direction ($\Pr(\hat{\beta} > 0) = 0.92$), with single-handed gestures having less pronounced directional differences than two-handed gestures.

In a separate analysis, we explored the role of social factors. For this, the values from the likeability scale and the AQ were added to created sum scores (one per participant). These scores were centered (to aid interpretation in the presence of interactions). A Bayesian logistic regression modeled the "Monday" versus "Friday" responses as a function of Gesture Direction, as well as the interaction between Gesture Direction and AQ and the interaction between Gesture Direction and Likeability.

For these social measures, we are specifically interested in the interaction effects with Gesture Direction and will only discuss these in the following. The AQ * Gesture Direction interaction was estimated to have a negative coefficient ($\hat{\beta} = -0.21$), indicating that participants with higher AQ (more autistic traits) responded less with "Friday" in response to forward movements. However, the 95% credible interval for this effect was very wide and included zero [-0.57,0.16]. The posterior probability of this effect being below zero was only $Pr(\hat{\beta} < 0) = 0.86$. For the likeability scale, more reliable results were obtained. Although the 95% of the coefficient ($\hat{\beta} = 0.24$) barely included zero [-0.04,0.52], compared to the effect of AQ, there was a stronger posterior probability associated with the effect being above zero, $Pr(\hat{\beta}) = 0.95$.

To get an idea of the strength of these results, it is easiest to discuss them in terms of a median split: participants with high likeability of the speaker reported

"Friday" 84% of the time for forward gestures, whereas participants with low likeability reported "Friday" only 75% of the time for these gestures. The effect was even more pronounced for backward gestures, where participants with high likeability of the speaker responded "Monday" 50% of the time, but participants with low likeability of the speaker only 25% of the time. These results are in line with the idea that if participants like the speaker more, they are more prone to assuming his perspective.

### 3.3. Discussion

Overall, Experiment 2 succeeded in replicating the Gesture Direction effect reported in Jamalian and Tversky (2012) and Lewis and Stickles (2017). While we recognize that there were some minor cosmetic improvements to the experiment that could have also been involved in producing differences with respect to Experiment 1, the Gesture Direction effect was most likely affected by the changes in hand shape. This supports our idea that the palm-up open hand gesture in Experiment 1 had alternative explanations. We found that there was little difference between the "flat" and the "push" hand shape, and there was equally relatively little difference between single-handed gestures and two-handed gestures, even though the latter were arguably more prominent.

There was some weak indication of social effects in the hypothesized direction: People who self-reported to have more (social) autistic traits were slightly less likely to assume the speaker's perspective. This effect was more pronounced for

the likeability scale than for the AQ scale. Participants who liked the speaker more were more likely to assume his perspective. In all future experiments, we will include the likeability scale to see whether we can accumulate evidence for the moderating role of speaker likeability.

The slight detour of Experiment 2 resulted in us finding a set of gestures that allow us to proceed with the original research question, investigating what happens when the temporal perspectives communicated via speech differ from those communicated via gesture. In the following, we will use the gesture that produced the most consistent effects (numerically), which was the two-handed flat-palm gesture. We now proceed with a replication of Experiment 1 ("backward" versus "forward" language), but with the new gestures and the added likeability scale.

## 4. Experiment 3

### 4.1. Methods

The methods were exactly as in Experiment 2, except for that this time, we re-included the "forward" and "backward" contrast from Experiment 1 (Language Direction). In addition, we only used one of the gestures explored in Experiment 2, namely, the two-handed flat-palm gesture, also used by Jamalian and Tversky (2012).

After noticing the large number of exclusions performed in Experiment 2, we decided to increase our sample size further so that the base sample (before exclusion) would at least be 300 participants. In the end, we collected data from a

total of 319 participants on Amazon Mechanical Turk for 0.40 USD reimbursement. After excluding non-native speakers, people who reported video issues, and people who did not respond either "Monday" or "Friday," we ended up with a total of 232 participants (35% exclusion). Of these, 125 were male, and 74 were female, with an average age of 34 (range 18 to 72).

### 4.2. Results

The proportions of "Monday" and "Friday" responses broken down by Gesture Direction and Language Direction are shown in Figure 4. Overall, participants chose "Monday" more often (69%) than "Friday" (31%). The results showed a strong effect of Language Direction, in the same direction as observed in Experiment 1, with noticeably more "Monday" responses ($N$ = 106, 90%) than "Friday" responses ($N$ = 12, 10%) for the "backward" adverb. In contrast, the "forward" adverb showed similar "Friday" ($N$ = 60, 53%) and "Monday" ($N$ = 54, 47%) responses. The Bayesian logistic regression indicated a strong Language Direction effect ($\hat{\beta} = 2.43$) for which the 95% credible interval was far away from zero, [1.71, 3.20]. In fact, not a single posterior sample was below zero, $Pr(\hat{\beta} > 0) = 1.0$.

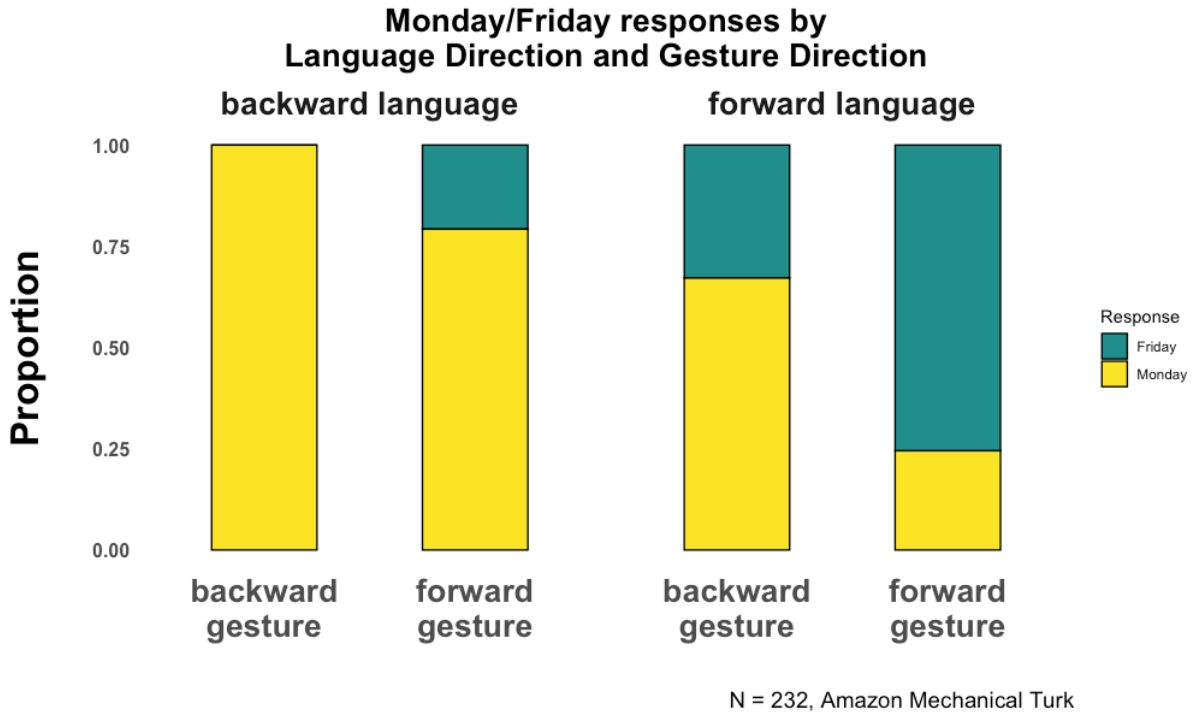


**Figure 4.** Results of Experiment 3 show a strong effect of Language Direction ("backward" language many more "Monday" responses), as well as a consistent effect of Gesture Direction (more "Friday" responses for forward movements)

In addition, we replicated the Gesture Direction effect observed in Experiment 2, with participants choosing "Monday" ($N$ = 101, 83%) much more often than "Friday" ($N$ = 20, 17%) when the gesture was moving backward. In contrast, participants chose "Monday" ($N$ = 59, 52%) about as often as "Friday" ($N$ = 52, 47%) when viewing forward-moving gestures. The logistic regression indicated a reliable Gesture Direction effect ($\hat{\beta} = 1.84$), with a credible interval far away from zero, [1.13, 2.58]. Again, not a single posterior sample was below zero, $Pr(\hat{\beta} > 0) = 1.0$. In terms of the overall magnitude, the Language Direction effect was 32% larger than the Gesture Direction effect. Comparisons of posterior samples of the respective

coefficients suggest that the posterior probability for the Language Direction effect being larger than the Gesture Direction effect was $Pr(\hat{\beta}_1 > \hat{\beta}_2) = 0.93$. There was no reliable Language Direction * Gesture Direction interaction ($\hat{\beta} = -0.23$, [-1.51, 0.99]).

Again, we found evidence for a weak interaction of Gesture Direction and likeability ($\hat{\beta} = 0.23$, [-0.01, 0.48]), with a high posterior probability of the interaction coefficient being above zero, $Pr(\hat{\beta} > 0) = 0.97$.

### 4.3. Discussion

Experiment 3 found Language Direction and Gesture Direction to have independent additive rather than multiplicative effects. That is, the forward/backward movement contrast did not differ reliably as a function of using "forward" or "backward" language. Overall, we found the Language Direction effect to be much stronger than the Gesture Direction Effect.

In addition, we found another weak but reliable interaction between perceived likeability of the speaker and the Gesture Direction effect, with participants who liked the speaker more strongly also showing stronger Gesture Direction effects.

## 5. Experiment 4

### 5.1. Methods

The lack of an interaction between the directions implied by adverbs and gesture seems to suggest some independence of gesture from the concomitant speech.

However, to really explore the dependence of gesture on the spatial adverbials, we need to conduct an experiment where gesture occurs without spatial adverbials. In Experiment 4, we asked, "Next Wednesday's meeting has been moved by two days, what day is it on now?" (without adverb), as well as, "Next Wednesday's meeting has been changed by two days, what day is it on now?" (without any indication of movement). The same gestures (two-handed flat palm) from Experiment 3 was used again, with forward and backward direction.

Overall, we had a total of 322 participants, of which 82 had to be excluded due to being non-native speakers, having experienced video issues, or not responding "Monday" or "Friday" (25% exclusion). The final sample consisted of 122 male participants, 114 female participants, and 4 others, with an average age of 37 (range 18 to 75).

**5.2. Results**

Figure 5 shows the proportion of "Monday" and "Friday" responses as a function of Language ("moved by" versus "changed by") and Gesture Direction. Overall, there was a disproportionate amount of "Friday" responses ($N$ = 191, 80%) compared to "Monday" responses ($N$ = 49, 20%).

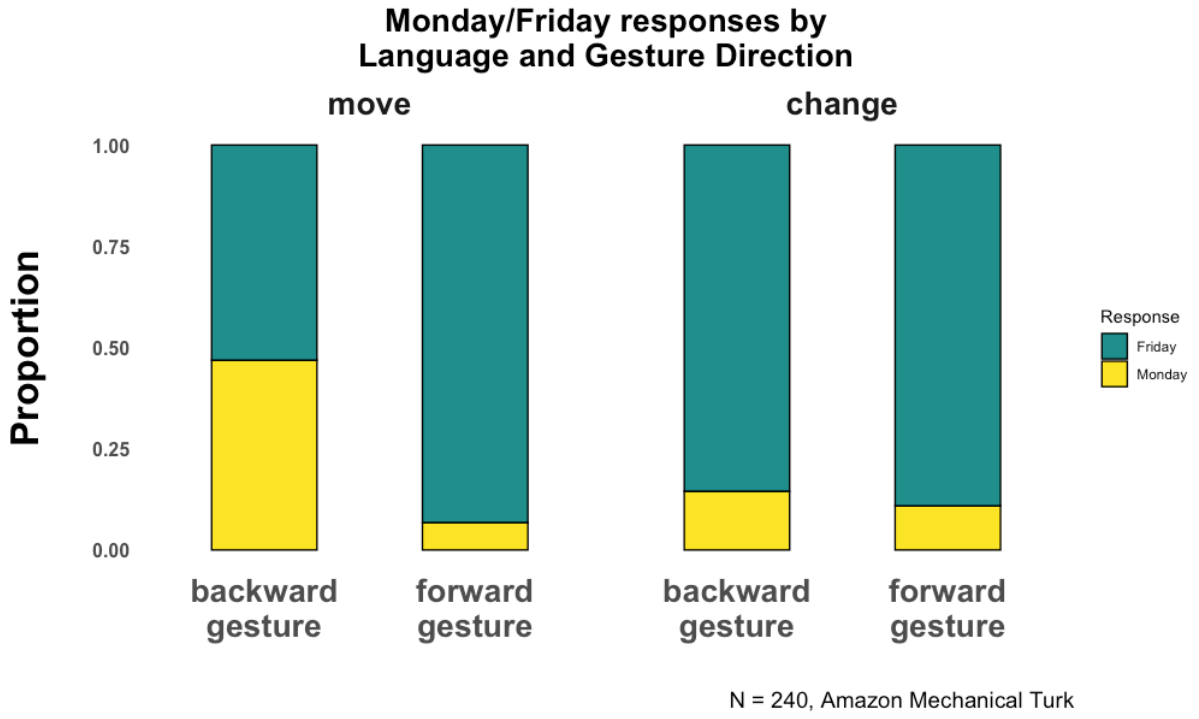


**Figure 5.** Results of Experiment 4 show an effect of Gesture Direction only for "moved forward two days," not for "changed by two days"

There was a relatively small main effect of Language: "changed by" had relatively more "Friday" responses ($N$ = 102, 87%) than "Monday" responses ($N$ = 15, 13%). In contrast, "moved by" had a slightly lower proportion of "Friday" responses ($N$ = 89, 72%), as well as a slightly higher proportion of "Monday" responses ($N$ = 34, 28%). The log odd coefficient of this effect ($\hat{\beta} = 0.67$) had a credible interval that barely included zero, [-0.02,1.38] and a relatively high posterior probability of being above zero, $Pr(\hat{\beta} > 0) = 0.97$.

There also was a main effect of Gesture ($\hat{\beta} = 0.19$) with a credible interval far away from zero [0.6, 2.02], $Pr(\hat{\beta} > 0) = 0.99$. Overall, forward movements lead to a very high number of "Friday" responses ($N$ = 104, 91%) compared to "Monday"

responses ($N$ = 10, 9%). In contrast, backward movements had slightly fewer "Friday" responses ($N$ = 87, 69%) and slightly more "Monday" responses ($N$ = 39, 31%).

By far the strongest pattern in this data was a highly reliable Gesture Direction * Language interaction ($\hat{\beta} = -1.43, [-2.64, -0.23]$), with a high posterior probability of being below zero $Pr(\hat{\beta} < 0) = 0.99$. The direction of this interaction is as follows: the forward/backward gestural contrast only played a strong role for the "moved by" question, not for the "changed by" question.

Finally, we found a weak interaction between Gesture Direction and likeability ($\hat{\beta} = 0.22, [-0.07, 0.52]$), with a posterior probability $Pr(\hat{\beta} > 0) = 0.93$.

**5.3. Discussion**

Results of Experiment 4 showed that Gesture Direction only produced a difference in "Monday" versus "Friday" responses when the language implied any sense of movement ("moved by"), not when the question merely referenced that the meeting has been "changed by" two days. In addition, we found more (weak) evidence for the effect of likeability.

A somewhat perplexing result is the overall high number of "Friday" responses in this task, compared to previous experiments. In fact, in Experiments 1 and 3, we observed more "Monday" responses (71% and 69% respectively), whereas in Experiment 2 and 4, we observed more "Friday" responses (56% and 80% respectively). What explains these differences? To investigate this, we conducted an

analysis across the different experiments, using information we had about the day of the week the study was collected. In addition, we used this conjoined analysis to study the reliability of the likeability effect across experiments.

## 6. Meta-analysis

### 6.1. Effect of day of the week

In this section, we explored a potential extraneous factor that may help to explain the differences in the overall proportion of "Monday" and "Friday" responses (regardless of any condition manipulations). We only do this descriptively as we did not predict these patterns and therefore do not want to perform confirmatory statistics on them.

For their original study, McGlone and Harding (1998) only conducted the experiment on a Wednesday. Experiment 1 was run continuously on several days of the week, including Monday, Wednesday, Thursday, Saturday, and Sunday. Experiment 2 was all run on a single day, Thursday. Experiments 3 and 4 were run at the same time on two consecutive days, Friday and Saturday. Figure 6 shows the proportion of "Monday" and "Friday" responses as a function of the day of the week. This shows clearly that conducting the experiment on a Monday elicits more "Monday" responses (77%). Similarly, the highest proportion of "Friday" responses (58%) was obtained when the experiment was conducted on a Friday. It is also noticeable that days closer to Monday and Friday exhibit progressively more "Monday" or "Friday" responses, respectively.

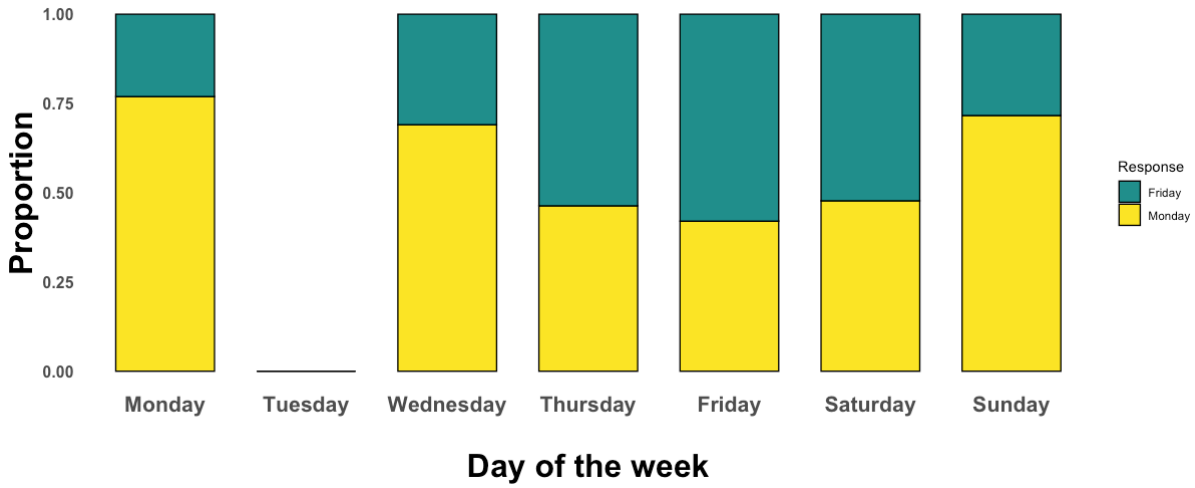


**Figure 6.** Results of Experiment 4 show an effect of Gesture Direction only for "moved forward two days," not for "changed by two days"

These differences of the day that the experiment was conducted may explain the difference between Experiment 1 and Experiment 2, since Experiment 1 (many more "Monday" responses) was conducted on several days of the week, including Monday, but Experiment 2 (many more "Friday" responses) was conducted on a Thursday, much closer to Friday. However, the day of the week differences cannot explain the stark difference in the proportion of "Monday" and "Friday" responses between Experiments 3 and 4, which were run at exactly the same time (these two experiments were released on Amazon Mechanical Turk together). This suggests that the difference between these two experiments has to do with the stimuli. One possible explanation is that in the absence of any directional information (no

mentioning of "forward" or "backward"), people assume a future oriented perspective by default.

### 6.2. Effect of likeability across experiments

In Experiments 2, 3 and 4, we observed weak but reliable interactions between likeability and Gesture Direction. However, for each of these experiments, the 95% credible interval covered zero. Thus, to assess the reliability of this finding, we wanted to accumulate the evidence for the likeability interaction effect across experiments, running a separate Bayesian logistic regression model that combined the data from all studies. This analysis revealed that across experiments, there is much evidence for the interaction between Likeability and Gesture Direction (see Figure 7), with the overall coefficient being positive ($\hat{\beta} = 0.19$, [0.05, 0.32]) and having a high posterior probability of being above zero $Pr(\hat{\beta} > 0) = 0.997$.

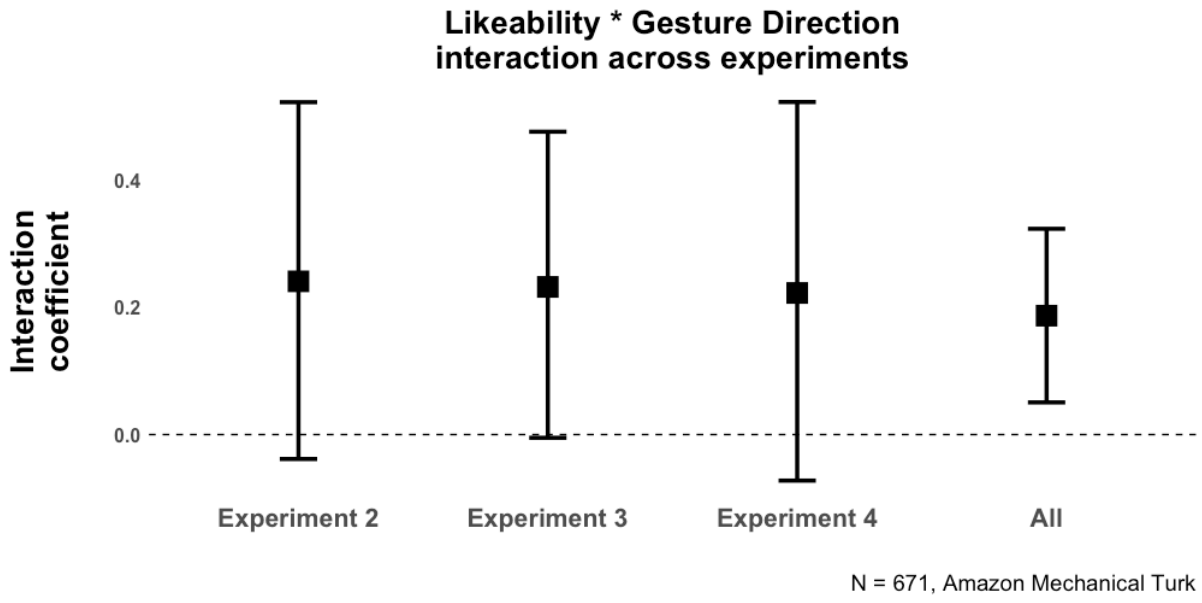

**Figure 7.** Likeability interaction effect across experiments 2 to 4 (likeability was not included in Experiment 1), as well as the coefficient of a combined analysis that accumulates evidence across experiments; error bars indicate 95% Bayesian credible intervals

It is also worth pointing out that the distribution on the likeability scale is dominated by one particular value, 12 (the median), which was the result for a total of 272 participants (40%), even though there is quite a range of possible responses. These seem to be participants who have just picked a middle-range response (slightly leaning to the positive side) for all four likeability questions. If we exclude the median and look at the descriptive statistics of the responses above and below the median, a clear picture emerges: Participants who reported low likeability of the speaker showed less strong results. They interpreted backward gestures to indicate "Monday" only 54% of the time, and forward gestures to indicate "Friday" only 68% of the time. In contrast, people who reported high likeability of the speaker, thought that a backward movement indicated "Monday" 70% of the time, and a forward movement indicated "Friday" 78% of the time—both of which are interpretations of the Next Wednesday's meeting question that are in line with the speaker's perspective.

It is possible that the likeability effect has an alternative explanation, namely, it could be that participants who did not like the speaker paid less attention to the gestures. We find this to be unlikely because our post-experiment debriefing

questions revealed that across experiments, almost all participants reported to have noticed the gestures in the video, 92% in Experiment 1, 79% in Experiment 2, 86% in Experiment 3, ad 85% in Experiment 4. However, it was indeed the case that participants who said that they did not notice the gesture at all ("Definitely not") also had slightly lower values on the likeability scale ($M$ = 10.2) than participants who reported "Definitely yes" ($M$ = 12.0). This suggests that our analysis of the likeability effect should control for this factor.

To rule out that the likeability effect may be due to differences in gesture awareness, we fitted another logistic regression model on "Monday" versus "Friday" responses that included the interaction between Gesture Direction and Likeability, as well as the interaction between Gesture Direction and Noticing (added as an ordinal factor, from "Definitely not" to "Definitely yes"). This model revealed almost exactly the same Gesture Direction * likeability interaction coefficient ($\hat{\beta} = 0.20$, [0.06, 0.42]) with a similarly high posterior probability of being above zero $Pr(\hat{\beta} > 0) = 0.997$. This suggests that the interaction between likeability on Gesture Direction is genuine and not confounded with gesture awareness.

**7. General discussion**

Overall, we have advanced the study of gesture-speech relationships, as well as the study of spatial conceptualizations of time in several different ways. First, our results show that in our task, co-speech gesture *cannot* carry the mental time line all by itself, at least in the context of the Next Wednesday's meeting question.

Specifically, the effect of Gesture Direction disappears if the linguistic context does not emphasize any sort of motion. Moreover, the effect of Gesture Direction is weaker than similar linguistic manipulations ("forward/backward"). This is even though we used very salient two-handed gestures in Experiments 3 and 4.

Of course, spatial orientation along the sagittal axis does matter when it comes to processing time more generally (e.g., Hartmann & Mast, 2012; Miles et al., 2010); thus, our results in no way speak against the idea of a time line. However, the results do suggest than when both speech and gesture encode information about movement along the mental time line, it is the linguistic information that takes precedence. Gesture can carry the mental time line, but only when the linguistic conditions are right.

The comparison between Experiment 1 and all following experiments furthermore suggests that the Gesture Direction effect depends on hand shape. If the hand assumes a palm-up open hand configuration, participants do not seem to think that the forward/backward gestures represent movement along the time line. This is consistent with the literature on the palm-up open hand gesture (Cooperrider et al., 2018; Kendon, 2004; Mittelberg, 2017; Müller, 2004; Parrill, 2008). While this gesture has been reported to have a complex web of potential meanings, depending on the context, many of these meanings have a "presenting" function. This suggests that when participants saw the gesture in our Experiment 1, they may have thought that the speaker merely wanted to show them some general information, rather than tell them something specifically about the passage of time.

On top of that, we obtained a new result which suggests that social factors influence perspective taking in gesture perception. Specifically, we observed stronger effects of the gestural manipulation when participants liked the speaker more. The most likely explanation for this is that in this situation, participants were more likely to assume the perspective of the speaker, which leads to more consistent forward = "Friday" and backward = "Monday" responses (seen from the speaker's perspective). Our post-hoc analysis (incorporating self-reported gesture awareness) ruled out a possible alternative explanation for this result, which is that the likeability effect may be due to gesture awareness. While it is the case that people liking the speaker were also more likely to notice his gestures, this did not "explain away" the influence of likeability.

The influence of likeability found in the present experiment will provide fruitful grounds for future studies. For example, a number of gesture researchers are interested in the concept of "viewpoint" in gesture (Parrill, 2010, 2012; Stec, 2012), e.g., when somebody is telling a story, they may adopt the viewpoint of the character mentioned in the story ("character viewpoint"), or the viewpoint of an external observer ("observer viewpoint"). This contrast has been shown to be relevant for studying children's development of narrative skills (Parrill et al., 2018) and the development of mathematical concepts (Gerofsky, 2010). Crucially, interpreting quick switches between character viewpoint and observer viewpoint gestures depends on one's ability to assume the perspective of the story teller. In addition, some of the same gestures (e.g., transversal gestures that move to the right

from the speaker's perspective) will have opposite interpretations depending on which perspective one assumes. The existence of a likeability effect in the present study suggests that it is useful to see whether people's interpretation of viewpoint gestures is influenced by likeability.

Our results also speak to the issue of perspective taking in communication more generally, and in the comprehension of gesture more specifically. Some studies find that addressees default to assuming an ego-centric perspective (de la Fuente et al., 2015; Pickering et al., 2012), and it has been argued that assuming an other-centric perspective may come at a cognitive cost (Epley et al., 2004; Horton & Keysar, 1996). However, other studies find that in certain tasks, people more readily assume an other-centric perspective even when this comes at a cognitive cost (Duran et al., 2011). These conflicting results have led Duran, Dale and Galati (2016) to propose a model where addressees probabilistically integrate multiple cues for whether they should or should not assume an ego-focused or other-focused perspective, including social cues (see also Galati & Avraamides, 2013). Within such a model, the likeability of the interlocutor can be seen as another factor that moderates people's propensity to assume other-focused perspectives.

It is a potential concern that we only tested videos with a frontal view of the speaker, as the willingness to assume the speaker's perspective may depend on whether one views the speaker face-to-face or from a side view. For this reason, Jamalian and Tversky (2012) asked the Next Wednesday's meeting question while standing next to the participant, so that the speaker and addressee had identical

points of view. The fact that the present experiments and Lewis and Stickles (2017) found overall more "Friday" responses for forward gestures despite using a face-to-face view suggests that it is not necessary to stand next to the participant, which ultimately suggests that there is a bias toward assuming an other-centric perspective in interpreting gestures in this task. Lewis and Stickles (2017) furthermore found that showing people videos from side-by-side view did not interact with how the forward/backward gestures were interpreted, which leads us to believe that the present results would not differ considerably if the video had been changed in this regard. However, it is quite possible that a side-by-side view may change the degree to which likeability moderates perspective taking, as the eyes, a key predictor of likeability (see L. S. Bohannon et al., 2013), may be less visible in this condition. This is a testable prediction that can be explored in future research.

Our results furthermore point to issues that need to be addressed in future research that uses the Next Wednesday's meeting question, or similar question formats. Specifically, we found quite varying differences in the overall proportion of "Monday" and "Friday" experiments across the four different experiments. At least part of these differences could be explained by looking at what day of the week the experiment has been conducted. Thus, future research needs to control for the day of the week. McGlone and Harding (1998) did in this in their original study, as did some subsequent studies (e.g., Duffy & Feist, 2017; Duffy et al., 2014) but, to the best of our knowledge, the day of the week has not been consistently controlled for in all studies on the Next Wednesday's meeting question. Keeping the day of the week

constant or at least factoring this into one's statistical analysis is especially important as the Next Wednesday's meeting question has started to be used in some clinical contexts (Elvevåg et al., 2011).

Nonetheless, even when taking day of the week into account, there were some unexplained differences in the overall rate of "Monday" and "Friday" responses across experiments. In particular, we currently have no compelling explanation for the difference between Experiment 3 (a lot of "Monday") and Experiment 4 (a lot of "Friday"), since these two experiments were conducted at the same time. One possible explanation is that the semantics of the specific linguistic items used in this study ("moved by" and "changed by") are future oriented when no spatial adverb ("forward" or "backward") is provided. Corpus linguistic analysis of these constructions could reveal in what contexts people are using these expressions, and whether these are used more often to refer to future events or not.

Another unexpected result is that we found that the adverb "forward" resulted in relatively more "Friday" responses than the adverb "backward," which resulted in more "Monday" responses. This was the case for both Experiment 1 and Experiment 3. These results were unexpected because they differ from what was previously reported by Feist and Duffy (2015), who observed a greater incidence of "Monday" responses for "forward". It could be that the presence of gestures has reversed the interpretation of these adverbs. This, however, is unlikely given that Experiment 1 failed to find an effect of Gesture Direction.

Another explanation has to do with the different study populations tested in Feist and Duffy (2015). Their participants were British English speakers residing in Newcastle-upon-Tyne; the participants in the present study were from the United States. It is possible that the precise semantics of "forward" and "backward" differ between these dialects. This is likely because dialectal differences in interpreting the Next Wednesday's meeting question have already been reported for other languages (German versus Swiss German, Stocker & Hartmann, 2019). In addition, prior research has shown that individual differences in future orientation, lifestyle choices, personality factors, and situational factors can influence responses to the Next Wednesday's meeting question (Duffy et al., 2014; Duffy & Feist, 2014), as well as one's "temporal focus" (whether one thinks the future is toward one's front or toward one's back) (De La Fuente et al., 2014; Li, 2018; Li & Cao, 2017, 2018). The participants in Feist and Duffy (2015) were university administrators, which is quite a different population than Amazon Mechanical Turk users in terms of profession and presumably also other social and psychological variables tied to the individual. Thus, it is possible that population differences explain the opposing results reported here and in Feist and Duffy (2015). Future research should also collect information on people's temporal focus, as was done in De La Fuente et al. (2014), to assess the extent to which it influences responses to the Next Wednesday's meeting question.

Overall, our results provide more evidence for the idea that people's thought about time is connected to thinking about space. Time and space are connected in language (Clark, 1973; Haspelmath, 1997; Lakoff & Johnson, 1980, 1999; Moore, 2014)

and thought (Boroditsky & Ramscar, 2002; Casasanto & Boroditsky, 2008), and differences in spatial terms ("forward" or "backward") or movement verbs ("moved by" versus "changed by") correspondingly influence people's reasoning about time. The experiments provided here furthermore attest to the psychological reality of the distinction between "Moving Time" and "Moving Ego" metaphors (Boroditsky & Ramscar, 2002; McGlone & Harding, 1998; Núñez et al., 2006), two perspectives that can be changed by gesture. However, although we have shown that gesture *can* influence temporal reasoning, thus replicating Jamalian and Tversky (2012) and Lewis and Stickles (2017), we have also shown that the effect of gesture is limited. The influence of gesture dwarfs in comparison to the influence of language even though we used highly salient gestures, and when the language of a message does not reference any movement, gesture ceases to influence the mental time line.

**Acknowledgments**

We thank Ellen Schepp-Winter, Mark Mallabone, and two anonymous reviewers for useful feedback and suggestions.

**References**

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202–212.

Bender, A., & Beller, S. (2014). Mapping spatial frames of reference onto time: A review of theoretical accounts and empirical findings. *Cognition*, *132*(3), 342–382. https://doi.org/10.1016/j.cognition.2014.03.016

Bohannon, J. (2011). Social science for pennies. *Science*, *334*(6054), 307.

Bohannon, L. S., Herbert, A. M., Pelz, J. B., & Rantanen, E. M. (2013). Eye contact and video-mediated communication: A review. *Displays*, *34*(2), 177–185.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2257–2273. https://doi.org/10.1016/j.neubiorev.2012.08.007

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, *13*(2), 185–189.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593.

Casasanto, D., & Jasmin, K. (2012). The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics*, *23*(4), 643 – 674. https://doi.org/10.1515/cog-2012-0020

Clark, H. H. (1973). Space, time, semantics, and the child. *Cognitive Development and Acquisition of Language*, 27–63.

Cooperrider, K., Abner, N., & Goldin-Meadow, S. (2018). The palm-up puzzle: Meanings and origins of a widespread form in gesture and sign. *Front. Commun. 3: 23. Doi: 10.3389/Fcomm*.

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, *14*(2), 159–190.

de la Fuente, J., Casasanto, D., & Santiago, J. (2015). Observed actions affect body-specific associations between space and valence. *Acta Psychologica*, *156*, 32–36.

De La Fuente, J., Santiago, J., Román, A., Dumitrache, C., & Casasanto, D. (2014). When you think about it, your past is in front of you: How culture shapes spatial conceptions of time. *Psychological Science*, *25*(9), 1682–1690. https://doi.org/10.1177/0956797614534695

Duffy, S. E., & Evans, V. (2017). The top trumps of time: Factors motivating the resolution of temporal ambiguity. *Language and Cognition*, *9*(2), 293–315. https://doi.org/10.1017/langcog.2016.8

Duffy, S. E., & Feist, M. I. (2014). Individual differences in the interpretation of ambiguous statements about time. *Cognitive Linguistics*, *25*(1), 29–54. https://doi.org/doi.org/10.1515/cog-2013-0030

Duffy, S. E., & Feist, M. I. (2017). Power in time: The influence of power posing on metaphoric perspectives on time. *Language and Cognition*, *9*(4), 637–647. https://doi.org/10.1017/langcog.2016.33

Duffy, S. E., Feist, M. I., & McCarthy, S. (2014). Moving through time: The role of personality in three real-life contexts. *Cognitive Science*, *38*(8), 1662–1674. https://doi.org/doi.org/10.1111/cogs.12124

Duran, N., Dale, R., & Galati, A. (2016). Toward integrative dynamic models for adaptive perspective taking. *Topics in Cognitive Science*, *8*(4), 761–779.

Duran, N., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, *121*(1), 22–40.

Elvevåg, B., Helsen, K., De Hert, M., Sweers, K., & Storms, G. (2011). Metaphor interpretation and use: A window into semantics in schizophrenia. *Schizophrenia Research*, *133*(1–3), 205–211. https://doi.org/10.1016/j.schres.2011.07.009

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327.

Evans, V. (2013). *Language and time: A cognitive linguistics approach*. Cambridge University Press.

Feist, M. I., & Duffy, S. E. (2015). Moving beyond 'Next Wednesday': The interplay of lexical semantics and constructional meaning in an ambiguous metaphoric statement. *Cognitive Linguistics*, *26*(4), 633–656.

Galati, A., & Avraamides, M. N. (2013). Flexible spatial perspective-taking: Conversational partners weigh multiple cues in collaborative tasks. *Frontiers in Human Neuroscience*, *7*, 618.

Gerofsky, S. (2010). Mathematical learning and gesture: Character viewpoint and observer viewpoint in students' gestured graphs of functions. *Gesture*, *10*(2), 321–343. https://doi.org/10.1075/gest.10.2-3.10ger

Goldin, G., & Darlow, A. (2013). *TurkGate (Version 0.4. 0)[Software]*.

Goldin-Meadow, S. (2005). *Hearing gesture: How our hands help us think*. Harvard University Press.

Hartmann, M., & Mast, F. W. (2012). Moving along the mental time line influences the processing of future related words. *Consciousness and Cognition*, *21*(3), 1558–1562.

Haspelmath, M. (1997). *From space to time: Temporal adverbials in the world's languages*. Lincom Europa.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117.

Jamalian, A., & Tversky, B. (2012). Gestures alter thinking about time. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *34*.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language.' *Phil. Trans. R. Soc. B*, *369*(1651), 20130293.

Kranjec, A. (2006). Extending spatial frames of reference to temporal concepts. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (Vol. 28, pp. 447–452). Lawrence Erlbaum Associates.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic Books.

Lewis, T. N., & Stickles, E. (2017). Gestural modality and addressee perspective influence how we reason about time. *Cognitive Linguistics*, *28*(1), 45–76. https://doi.org/10.1515/cog-2015-0137

Li, H. (2018). A future-minded lark in the morning: The influence of time-of-day and chronotype on metaphorical associations between space and time. *Metaphor and Symbol*, *33*(1), 48–57. https://doi.org/10.1080/10926488.2018.1407995

Li, H., & Cao, Y. (2017). Personal attitudes toward time: The relationship between temporal focus, space-time mappings and real life experiences. *Scandinavian Journal of Psychology*, *58*(3), 193–198. https://doi.org/10.1111/sjop.12358

Li, H., & Cao, Y. (2018). The hope of the future: The experience of pregnancy influences women's implicit space–time mappings. *The Journal of Social Psychology*, *158*(2), 152–156. https://doi.org/10.1080/00224545.2017.1297289

Loermans, A. C., de Koning, B. B., & Krabbendam, L. (2019). Agency and time representation in English and Dutch speakers. *Journal of Language and Social Psychology*, *38*(3), 353–375. https://doi.org/10.1177/0261927X18824619

Matlock, T., Holmes, K. J., Srinivasan, M., & Ramscar, M. (2011). Even abstract motion influences the understanding of time. *Metaphor and Symbol*, *26*(4), 260–271. https://doi.org/10.1080/10926488.2011.609065

McGlone, M. S., & Harding, J. L. (1998). Back (or forward?) to the future: The role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1211–1223.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Miles, L. K., Karpinska, K., Lumsden, J., & Macrae, C. N. (2010). The meandering mind: Vection and mental time travel. *PLoS One*, *5*(5).

Mittelberg, I. (2017). Multimodal existential constructions in German: Manual actions of giving as experiential substrate for grammatical and gestural patterns. *Linguistics Vanguard*, *3*(s1). https://doi.org/10.1515/lingvan-2016-0047

Moore, K. E. (2006). Space-to-time mappings and temporal concepts. *Cognitive Linguistics*, *17*(2), 199–244. https://doi.org/10.1515/COG.2006.005

Moore, K. E. (2014). *The spatial language of time: Metaphor, metonymy, and frames of reference*. John Benjamins.

Müller, C. (2004). Forms and uses of the Palm Up Open Hand: A case of a gesture family. In C. Müller & R. Posner (Eds.), *The semantics and pragmatics of everyday gestures* (pp. 233–256). Weidler.

Müller, C. (2009). *Metaphors dead and alive, sleeping and waking: A dynamic view*. University of Chicago Press.

Núñez, R., Cooperrider, K., Doan, D., & Wassmann, J. (2012). Contours of time: Topographic construals of past, present, and future in the Yupno valley of Papua New Guinea. *Cognition*, *124*(1), 25–35. https://doi.org/10.1016/j.cognition.2012.03.007

Núñez, R., Motz, B. A., & Teuscher, U. (2006). Time after time: The psychological reality of the ego-and time-reference-point distinction in metaphorical construals of time. *Metaphor and Symbol*, *21*(3), 133–146. https://doi.org/10.1207/s15327868ms2103_1

Núñez, R., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*, *30*(3), 401–450.

Ouellet, M., Santiago, J., Israeli, Z., & Gabay, S. (2010). Is the future the right time? *Experimental Psychology*, *57*, 308–314. https://doi.org/10.1027/1618-3169/a000036

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184–188. https://doi.org/10.1177/0963721414531598

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419.

Parrill, F. (2008). Form, meaning, and convention: A comparison of a metaphoric gesture with an emblem. In A. Cienki & C. Müller (Eds.), *Metaphor and gesture* (pp. 195–217). John Benjamins.

Parrill, F. (2010). Viewpoint in speech–gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes*, *25*(5), 650–668. https://doi.org/10.1080/01690960903424248

Parrill, F. (2012). Interactions between discourse status and viewpoint in co-speech gesture. In B. Dancygier & E. Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 97–112). Cambridge University Press.

Parrill, F., Lavanty, B., Bennett, A., Klco, A., & Demir-Lira, O. E. (2018). The relationship between character viewpoint gesture and narrative structure in children. *Language and Cognition*, *10*(3), 408–434. https://doi.org/10.1017/langcog.2018.9

Pickering, M. J., McLean, J. F., & Gambi, C. (2012). Do addressees adopt the perspective of the speaker? *Acta Psychologica*, *141*(2), 261–269.

Ramscar, M., Matlock, T., & Dye, M. (2010). Running down the clock: The role of expectation in our understanding of time and motion. *Language and Cognitive Processes*, *25*(5), 589–615. https://doi.org/10.1080/01690960903381166

Reysen, S. (2005). Construction of a new scale: The Reysen likability scale. *Social Behavior and Personality: An International Journal*, *33*(2), 201–208. https://doi.org/10.2224/sbp.2005.33.2.201

Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior*, *43*, 304–307.

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, *43*(1), 155–167.

Stec, K. (2012). Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture*, *12*(3), 327–360. https://doi.org/doi.org/10.1075/gest.12.3.03ste

Stickles, E., & Lewis, T. N. (2018). Wednesday's meeting really is on Friday: A meta-analysis and evaluation of ambiguous spatiotemporal language. *Cognitive Science*, *42*(3), 1015–1025. https://doi.org/doi.org/10.1111/cogs.12559

Stocker, K., & Hartmann, M. (2019). "Next Wednesday's meeting has been moved forward two days": The time-perspective question is ambiguous in Swiss German, but not in standard German. *Swiss Journal of Psychology*, *78*(1–2), 61. https://doi.org/10.1024/1421-0185/a000220

Streeck, J. (2009). *Gesturecraft*. John Benjamins.

Walker, E., & Cooperrider, K. (2016). The continuity of metaphor: Evidence from temporal gestures. *Cognitive Science*, *40*(2), 481–495.

Weger, U. W., & Pratt, J. (2008). Time flies like an arrow: Space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review*, *15*(2), 426–430. https://doi.org/10.3758/PBR.15.2.426

Winter, B., & Duffy, S. E. (2020). *Can co-speech gestures carry the mental time line?* https://doi.org/10.17605/OSF.IO/VC8BW

Winter, B., Marghetis, T., & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex*, *64*, 209–224.